

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12P 21/00, C07K 1/04 // 1/13		A1	(11) International Publication Number: WO 98/21353 (43) International Publication Date: 22 May 1998 (22.05.98)
(21) International Application Number: PCT/SE97/01805 (22) International Filing Date: 29 October 1997 (29.10.97)		(81) Designated States: JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(30) Priority Data: 9604108-2 8 November 1996 (08.11.96) SE		Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(71)(72) Applicants and Inventors: EHRENBERG, Måns [SE/SE]; S:t Olofsgatan 24, S-751 21 Uppsala (SE). PAVLOV, Michail [—/SE]; Kvärnbäcksgatan 1, S-754 20 Uppsala (SE).			
(74) Agents: BERGANDER, Håkan et al.; Pharmacia & Upjohn AB, Patent Department, S-751 82 Uppsala (SE).			

(54) Title: IN VITRO PRODUCTION OF PROTEINS BY TRANSLATION OF mRNA

(57) Abstract

An in vitro method for production of a protein (polypeptide), characterised in that the mRNA encoding the protein is translated when bound to a solid phase. By allowing for a pause at a terminating stop codon and addition of release factor the final protein may be obtained in almost pure form. By allowing for pauses and restart at internal codons and replacement of the amino acid mixture used before a pausing with an amino acid mixture that differs with respect to labelling, region-labelled polypeptides/proteins can be accomplished. A labelled polypeptide/protein, characterised in that its amino acid sequence contains one or more regions of two or more labelled amino acid residues in sequence. In a subaspect, the polypeptide/protein has at least one amino acid that is occurring twice and this amino acid differs in labelling at at least two positions.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	MU	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

IN VITRO PRODUCTION OF PROTEINS BY TRANSLATION OF mRNA

5 TECHNICAL FIELD

The present invention concerns, firstly, solid phase translation in general of mRNA to give a protein (polypeptide), encoded by the mRNA, and secondly, as a subaspect, region specific labelling of one or more predetermined regions (part 10 sequences) of a polypeptide chain (protein) by solid phase translation of mRNA in vitro. The invention also encompasses novel region specific labelled proteins/polypeptides. The labelled polypeptides have their primary use in structural studies by NMR.

15

Technical background.

The inventive method generalises an earlier mRNA-analogue (poly(U)) column translation method (Belitsina et al, 1975; Belitsina and Spirin, 1979) refined by Baranov et al., (1979), used to obtain ribosome in pre- and post-translocational states. Poly(U) as used in this earlier technique is not mRNA since it does not contain all the elements necessary for normal translation (stop codon, SD-sequence, ribosome binding site etc). Homopolymers of amino acids are not proteins. In the context of 25 the invention homo polymers of amino acids are also excluded from the concept of polypeptides. In vitro translation has been described previously (e.g. Pavlov and Ehrenberg, 1996, and Ehrenberg et al, 1990).

NMR spectroscopy has over the past decade become a very 30 powerful method to determine structures of small proteins in solution (Bax, 1989; Schwabe et al, 1990; Härd et al, 1990; Baumann et al, 1993; van Tilborg et al, 1995). NMR has the intrinsic limitation that, as the studied proteins get larger, there is a drastic reduction in the resolution of their NMR 35 spectra (Bax, 1989). This drawback has been partially overcome by the application of various isotope labelling strategies (Muchmore

et al, 1989; Ramesh et al, 1994). At the same time, there is still a pronounced upper limit around 30 kD for the determination of protein structures at high resolution using NMR spectroscopy (Bax, 1989).

5 One may identify three major types of isotope labelling strategies for NMR studies. The first is "uniform" labelling, where all the different amino acids in a polypeptide are labelled with, e.g., ^{15}N , ^{14}C or ^2H isotopes. A combination of ^{15}N , ^{14}C or ^2H "uniform" labelling recently made it possible to determine the 10 structure in solution of such a large molecular complex as the trp repressor in complex with operator DNA (Zhang et al, 1994).

A second strategy is "selective" isotope labelling of proteins. This means that only a limited class of isotope labelled amino acids are built into the polypeptide, while the 15 other amino acids are unlabelled. Selectively labelled proteins are obtained from over-producing bacterial strains, which grow in media where one or several types of amino acids are isotope labelled. This strategy has become very useful for structural analysis with NMR (Muchmore et al, 1989; Ramesh et al, 1994 and 20 references therein).

A third strategy, which we denote "region specific" labelling, is more difficult to implement technically. This strategy means that labelled amino acids are incorporated only in one or more predetermined regions. One or more of all amino acid residues of 25 a given peptide region may be labelled, while amino acids located outside the region may be unlabelled. We judge this, third strategy as the potentially most powerful way to extend the range of NMR-spectroscopy beyond the 30 kD limit to larger protein structures and complexes.

30 One way to obtain region specific labelling of proteins is by chemical polypeptide synthesis (Boutillon et al, 1995). At present, this method can only be used for small proteins.

Objectives of the invention

A first objective is to provide an improved general method for in vitro translation enabling direct production of proteins in almost pure form.

5 A second objective is to provide a general method for region specific labelling of proteins based on in vitro translation as described e.g. by Pavlov and Ehrenberg, 1996, and Ehrenberg et al, 1990.

10 A third objective is to apply the solid support translation technology of the invention for implementing synthesis of region labelled proteins.

15 A fourth objective is to provide proteins that are isotope labelled at one or more predetermined regions.

15 **The invention**

These objectives can be accomplished by a method that contemplates translation of real mRNAs stably linked to a solid phase to give real proteins (polypeptides). These mRNAs contain in frame codons to be translated to an amino acid sequence. For 20 prokaryotes they also contain a Shine and Dalgarno sequence, a ribosomal binding site, an initiation codon and a termination codon, but one or several of these additional features may be deleted. For eucaryotes there is normally a cap structure at the 5'-end of the mRNA. One major advantage of this technique is that 25 the ribosomes can be stalled at the stop codon that signals that the protein is full length as long as release factor is not included in the translation mixture. Subsequently, all components and factors necessary for translation can be rinsed off the solid phase in a simple way. After this step the solid phase linked 30 mRNAs hold the ribosomes which hold the peptidyl-tRNAs that contain the protein of interest. Addition of the appropriate release factor (RF1 for UAA, UAG and RF2 for UAA or UGA) hydrolyses peptidyl-tRNA removing the protein of interest from the ribosomes that remain immobilized on the solid phase. Another 35 rinsing step elutes the protein of interest together with

catalytic amounts of RF1/2, making the final purification very simple.

Accordingly, the inventive method for production of proteins (polypeptides), is characterized in that mRNA encoding a protein of interest and bound to a solid phase is translated in vitro. In order to be able to obtain a highly purified form of the protein directly from the column, translation is preferably done in two steps: first with a translation mixture containing all components for translation to a stop codon but devoid of the appropriate release factor activity - removal of the translation mixture - addition/introduction of the appropriate release factor activity.

A complete translation mixture is normally in the form of an aqueous buffer solution and allows for translation of the complete mRNA of interest and release of the so expressed protein from the tRNA--ribosome--mRNA complex. The mixture thus contains all ingredients necessary for translation, i.e. ribosomes, amino acids, amino acyl tRNA synthetases, tRNAs, initiation factors, elongation factors, energy giving system, buffering substances etc. See for instance the experimental part and references cited therein. The exact composition will depend on the origin of the various enzymes and factors utilized. For an E. coli origin, for instance, specific components, such as fmet-tRNAtmet, may have to be included.

By manipulating the translation mixture, translation may be paused and restarted at predetermined positions of the mRNA. The translation mixture may be added step-wise to the mRNA to be translated, such as a first portion (mixture) comprising ribosomes, initiation factors and for a prokaryotic system fmet-tRNAtmet (for an eucaryotic system the initiator is met-tRNAmet) (initiation mix) followed by a second portion (mixture) comprising elongation factors, amino acyl tRNA synthetases, transfer RNAs, amino acids, energy giving system etc (translation mix). In the preferred case the translation mixture is added stepwise with mixes varying in composition enabling pausing and restarting of translation at predetermined and well-defined positions.

A pause in translation may be achieved by eliminating from a mixture the amino acyl-tRNA activity that reads the codon, where the translating ribosome shall stop. This may be achieved by removing the corresponding amino acid from the translation 5 mixture, preferably in conjunction with a defective tRNA synthetase activity for that amino acid. Alternatively, the pausing may be at an internal stop codon in the mRNA. Readthrough of the stop codon is achieved by introducing suppressor tRNA activity that is specific for the stop codon, by adding 10 suppressor tRNA charged with the correct amino acid at that position in the polypeptide sequence. Still another alternative for creating a pause at a defined codon is to use a mix which is lacking the isoacceptor tRNA activity for that codon. As indicated above it is often preferred to use two alternatives in 15 conjunction for pausing at a desired codon. Restart is accomplished by addition/introduction of the lacking activity (activities).

By arranging for pausing at two or more internal codons, it is possible to define internal regions that can be translated 20 specifically in the presence of a desired labelled amino acid mixture.

By running the translation before and after a pause with mixtures containing amino acids differing in labelling, region 25 specific labelled proteins (polypeptides) will be obtained. A specifically labelled region may contain two or more amino acid residues in sequence. The maximal length of a region is the total length of the protein (polypeptide) produced minus one or two amino acid residues. Specifically labelled regions of one amino acid residues may be better obtained by methods not belonging to 30 the invention.

The label may be of any type provided it is compatible with translation. This in most cases means that the label is an isotope of one or more of the basic elements normally occurring in native amino acids, i.e. an isotope of N, C, O, H and S, that 35 may be either radioactive (e.g. ^{3}H , ^{35}S , ^{14}C) or non-radioactive (e.g. ^{15}N , ^{13}C , ^{2}H). Isotope labelling includes cases where a

larger part of the protein (polypeptide) contains a "non-normal" isotope (e.g. ^2H (deuterium)) while the region of interest contains the "normal" isotope (e.g. ^1H (protons)). Labelling with isotopes means that the labelled amino acid carries more than 5 normal amounts of the isotope concerned.

In one preferred mode of the "labelling" aspect of the invention, the mRNA encoding the polypeptide to be produced is contacted with a first translation mix A that permits translation but is deficient in at least one of the above-mentioned 10 activities causing readthrough of a codon at which pausing is desired. In a subsequent step the nascent polypeptide ribosome-mRNA-solid phase complex is contacted with a second translation mix B that permits readthrough of the codon at which the 15 translation previously had stopped for mix A. Readthrough can be accomplished by including in mix B the activity/activities causing the pausing for mix A. In analogy with mix A, mix B can be deficient in a component specific for a predetermined amino acid encoding codon or for a stop codon (downstream the pausing position for mix A), which means that translation in conjunction 20 with mix B will be halted at this second predetermined codon. Optionally, mix B may be replaced with further translation mixes (C, D, and so on), such that a subsequent mix permits restarting 25 of translation from a predetermined codon at which the translation is paused for the closest preceding mix. By allowing the amino acid compositions for juxta-positioned translation mixes to differ with respect to labelling of one or more amino acids, region specific labelling may be achieved in any 30 preselected region/regions of interest in a protein. By providing release factor (RF1 and RF2) adapted to the stop codon used (UAA, UAG, UGA) the completed polypeptide will be released from the tRNA-ribosome-mRNA complex attached to the solid phase.

The temperature is selected according to rules known in the art, which normally means 10-40°C for systems derived from eubacteria or eucaryotes, bearing in mind that a too low 35 temperature will give a slow translation and a too high temperature will result in denaturation and inactivation of the

enzymes and other proteins involved. For systems derived from archaeabacteria considerably higher temperatures may be used. The pH conditions are also selected as outlined in the art for translation in vitro.

5 The solid phase used (support, carrier) are as a rule of the same type as those used as chromatographic adsorbents or supports for solid phase synthesis of oligo and poly nucleotides/peptides and encompass particulate as well as monolithic material, all of which in the preferred form should exhibit a hydrophilic contact
10 area towards the liquid medium used. The support may be porous or non-porous. The link between the solid support and the mRNA may be of any type as long as it can withstand the conditions applied during translation/initiation and is compatible with efficient translation. Preferably the link is of covalent nature, for
15 instance by attachment to the support via a terminal part, preferably a non-coding spacer sequence, of the mRNA, with preference for the 3'-end of the mRNA. Potential alternative ways of linking mRNA to supports involve attachment of affinity ligands to terminal parts of mRNA in combination with
20 complementary receptors attached to the support. Biotin-streptavidin and anti-hapten high affinity antibodies are interesting ligand-receptor pairs that can be used.

Solid phase bound mRNA may also be a consequence of transcribing solid phase bound DNA encoding the protein/
25 polypeptide of interest (e.g. cDNA). In this type of binding one may start amplifying target DNA by PCR using two primers one of which is biotinylated. The amplified biotinylated DNA is then transferred to a streptavidine coupled solid phase where it becomes firmly bound to the solid phase through a streptavidine-biotin complex. The DNA is then transcribed by an RNA polymerase
30 to prepare mRNA. mRNA is retained in complex with the RNA polymerase used and the DNA and can be translated in the same manner as for mRNA directly bound to the solid phase.

The main advantage with mRNA-linked solid supports is that a
35 nascent or finished polypeptide, still linked to peptidyl-tRNA ribosome and mRNA, can easily be separated from all other

components in a translation mix by simple washing. This makes switching between labelled and unlabelled amino acids easy, and gives almost pure protein in the final elution step.

After the complete translation and release of the polypeptide 5 from the ribosome, the polypeptide should be isolated and purified. This can be done by techniques well known in the art, bearing in mind that the use of solid supports carrying the mRNA to be translated is very advantageous for quickly obtaining a highly purified labelled polypeptide.

10 One aspect of the invention comprises a labelled polypeptide/ protein characterized in that its amino acid sequence contains one or more regions of two, three or more labelled amino acid residues in sequence. The label is preferably an isotope as described above. The two or more labelled amino acids that are in 15 sequence may be the same or different. By region is meant part sequences of two or more amino acids situated at chosen places in the sequence of the full length polypeptide/protein. The labelled part sequence may extend up to the full length protein although in order to be called a part sequence at least one or more amino 20 acid must be missing. In a subaspect of this aspect, at least one amino acid is occurring at least twice, with the provision that the labelling of such an amino acid differs in at least two positions.

25 The best mode of the invention at the filing date is presented in the Experimental Part.

EXPERIMENTAL PART

30 MATERIALS AND METHODS

General description of the technique

Synthesis of region labelled proteins involved the following consecutive steps:

35 a) An initiation mix, containing ribosomes, initiation factors and precharged fMet-tRNA^{fMet}, is applied to a

column with mRNAs 3'-covalently coupled to a solid support.

- b) A translation mix, containing an energy regenerating system (Jelenc and Kurland, 1979), elongation factors, tRNA synthetases, bulk tRNA and unlabelled amino acids, is applied to the column. The ribosomes, initiated in step (a), translate the column coupled mRNA to the first codon of the region of interest, where they pause.
- c) The first translation mix is washed out from the column and a second translation mix, containing labelled amino acids, is applied. The ribosomes translate to the last codon of the region of interest, where they pause again.
- d) The second translation mix is washed out and a third translation mix containing unlabelled amino acids, is added to the column and the ribosomes translate the mRNA to a STOP codon, where they pause.
- e) The third translation mix is washed out from the column, release factor is applied and almost pure region labelled protein is eluted.

20

Chemicals

Phosphoenol pyruvate, putrescine, spermidine, myokinase and non-radioactive amino acids were from Sigma (U.S.A.). Pyruvate kinase was from Boehringer (Germany). All radioactive chemicals were from Amersham (United Kingdom). Other chemicals of analytical grade were from Merck, USA. Adipic acid Hydrazide Agarose was from Pharmacia Biotech AB (Uppsala, Sweden).

General components used in the translation system.

30 Initiation factors were purified from overproducing strains according to Soffientini et al (1994) with minor modifications. Elongation factors Tu, Ts and G were purified according to Ehrenberg et al (1990). Bulk tRNA was prepared according to Kelmers et al (1971). Ribosomes were purified from *E. coli* strain 35 017 (Zengel et al, 1977) according to Jelenc (1980) and stored in polymix buffer (Jelenc and Kurland, 1979) at -80°C. tRNA^{fMet} was

isolated from bulk *E. coli* tRNA as described by Seno et al (1968). Acylation and formylation of tRNA^{fMet} were according to Dubnoff and Maitra (1971). [³H]Met-tRNA was purified to near 100% on BD-cellulose and the final degree of formylation was 80%.
5 Release factor RF1 was purified from an overproducing strain as for initiation factors

Crude synthetase mixtures (denoted RS30) were prepared from the S30 supernatant from *E. coli* cells. The supernatant was first applied to a DEAE-Sepharose CL 6B column in TM buffer [40 mM Tris 10 (pH 7.5); 5 mM MgCl₂; 2 mM DTE; 10 % glycerol; 100 mM PMSF; 0.02% NaN₃] and then eluted with a 0-500 mM NaCl gradient in the same buffer. The synthetase containing fractions starting from GlnRS and ending with LysRS, were pooled, precipitated with ammonium sulphate, dialysed against polymix buffer and stored at -20°C 15 after addition of glycerol (50% final concentration).

Four different synthetase mixtures were used:

- (a) mixture RS30WT, purified from MRE600 (wt) *E. coli* cells.
- (b) mixture RS30-ThrRS, purified from strain IBQC6111 (Corner et al, 1996, and references cited therein), with a 20 deficient ThRS.
- (c) mixture RS30-GlnRS, purified from strain IBPC420 (Corner et al, 1996, and references cited therein), with a deficient (temperature sensitive) GlnRS.
- (d) mixture RS30-RF1, purified from strain MRA8 (Zang et al, 25 1994) with a deficient release factor RF1.

Construction of a modified GR DBD gene and preparation of mRNA for GR DBD: An artificial gene, GR8UAG, for the DNA binding domain (DBD) of the human glucocorticoid repressor (GR) was 30 derived by PCR technique using the forward primer

(5') gaaatataatcgactcaactatagGGTTAACCTTAgtaAGGAGgTaaaaaaaa
ATGAAACTGTGCCTGGTatTGC

SEQ ID NO 1

35 and the reverse primer

(5') ttttttttttttttttttCTGCAGctgctactgTattCCTTTATTTTTCTTG
SEQ ID NO 2

The plasmid pT7GR8 (obtained from Anthony Wright) which contains
5 the coding sequence of GR DBD was used as a template (sequences
of the primers complementary to the plasmid sequences are
underlined). The resulting artificial gene contained a promotor
for T7 RNA polymerase, a ribosome binding site (RBS) with the
epsilon chain (Olins et al., 1989) and a strong Shine-Dalgarno
10 (SD) sequence (Gold et al., 1990):

15 derived from the forward primer. The gene also contained a poly(A) tail, derived from the reverse primer. In addition the original ochre stop codon UAA in the plasmid sequence of GR DBD was substituted with an amber stop codon (shown in bold face type 20 in the reverse primer sequence (SEQ ID NO 2)), read by release factor RF1 but not by RF2. The poly(A) tail simplifies isolation of full length mRNA transcripts (below).

The reaction mix (4 ml) for in vitro transcription contained 40 mM Tris (pH 7.9), 22 mM MgCl₂, 2 mM spermidine, 2 mM DTE; 4 mM each of ATP, UTP, CTP, and GTP, 150 pmol GR8UAG, 100 units 25 RNAGuard (Pharmacia Biotech AB) and 2000 units T7 RNA polymerase. After 3 h incubation at 37°C the reaction was stopped by addition of NaCl and EDTA to the final concentrations of 0.5 M and 20 mM, respectively. The mix was subsequently cooled, diluted with 30 buffer A [10 mM Tris (pH 7.5), 500 mM NaCl, 1 mM EDTA] and applied to a 2 ml oligo(dT) cellulose column (Pharmacia Biotech AB). Non-bound material was removed from the column by washing with buffer A and GR8UAG mRNA was eluted with buffer A lacking NaCl.

35 Preparation of mRNA resin and column translation. 1.5 nmol of Gr8UAG mRNA was incubated for 1 h in the dark at room temperature

in 1.0 ml of 100 mM NaOAc (pH 5.1) buffer with 10 mM freshly prepared sodium metaperiodate (NaIO₄). The oxidised mRNA was ethanol precipitated, the pellet was washed with 80 % ethanol and re-dissolved in 1.0 ml of 100 mM acetate buffer (pH 5.1).

5 Finally, the pellet was dissolved in 1 ml of 100 mM H₂O and passed through a NAPD10 column (Pharmacia Biotech AB) equilibrated with H₂O to remove traces of NaIO₄. Adipic Acid Hydrazide Agarose (Pharmacia Biotech AB) was washed extensively with water. The resin (4 ml) was then resuspended in 10 ml of 200 10 mM sodium acetate (pH 5.1) and 1.5 ml of oxidised GR8UAG mRNA from the NAPD10 column was subsequently added. The coupling reaction (spontaneous coupling) was allowed to proceed for 20 h at 6°C with mild shaking. Then, an equal volume of pure glycerol was added to the resin (now denoted mGRU8UAG), which subsequently 15 was stored at -20°C. From the OD of the column wash, it was estimated that more than 90 % of the applied mRNA was coupled to the resin.

To prepare a column for translation, 2 ml of mGRU8UAG resin was transferred into a Poly Prep Chromatography column (BioRad, USA). 20 The column was washed with 8 ml water, then with 10 ml of 2 M KCl and finally with 8 ml water, to remove glycerol and unbound material. Before each translation experiment, the column was pre-equilibrated with polymix buffer (see below) containing 1 mM GTP. The OD was monitored to detect mRNA dissociation. The final bed 25 volume of the column was 0.4 ml and the calculated density of mRNA was 0.35 nmol/ml. The total amount of mRNA in the column was 140 pmol.

Mixes for translation experiments.

30 All mixes described were prepared on ice in polymix buffer (Jelenc and Kurland, 1979; Wagner et al, 1982). Polymix buffer contained, at working strength, 5 mM magnesium acetate, 5 mM potassium phosphate (pH 7.3), 95 mM potassium chloride, 5 mM ammonium chloride, 0.5 mM calcium chloride, 1 mM spermidine, 8 mM 35 putrescine and 1 mM 1,4-dithioerythritol (DTE).

Initiation mixes. These contained 1 mM GTP, 0.5 unit/ml RNA guard (Pharmacia Biotech AB) as well as different amounts of 70S ribosomes, initiation factors IF1, IF2, IF3 and fMet-tRNA^{fMet} as specified for each experiment, but always with the molecular ratios 2(Rib):3(IF1):2(IF2):3(IF3):3(fMet-tRNA^{fMet}). Different translation mixes were used to prepare a region labelled GR DBD. They all contained 45 µM EF-Tu, approximately 60 µM of bulk tRNA (estimated from OD at 260 nm), 4.5 µM EF-G, 2 µM EF-Ts, 1 mM ATP, 10 mM phosphoenolpyruvate, 0.8 mM GTP, 0.2 mM CTP, and also (per 100 ml) 5 µg pyruvatekinase (PK), 0.3 µg myokinase (MK), 0.2 OU (at 280 nm) of the corresponding RS30 synthetase mix. Concentrations of added amino acids were as specified in each experiment.

Translation mixes for column translations. To synthesise in the column a region labelled domain of glucocorticoid receptor (GR DBD) three translation mixtures were prepared. Mix A (TMA) contained RS30 synthetase mix with defective ThrRS and also all unlabelled amino acids except Thr and Trp at 0.1 mM concentration. Mix B (TMB) contained RS30 synthetase mix with defective GlnRS and nine ¹⁴C labelled amino acids: A, E, F, G, K, R, S, T and V with a specific activity of approximately 1500 dpm/pmol at approximately 5 µM concentration each. Mix C (TMC) contained RS30 synthetase mix with defective RF1 and all unlabelled amino acids except Trp at 0.1 mM concentration. All 25 translation mixtures were pre-incubated 10 min at 39°C before they were applied to the mRNA-column.

Gel electrophoresis and data acquisition.

Samples for SDS-gel electrophoresis were incubated for 30 min 30 at 42°C in 0.2 M freshly prepared NaOH to hydrolyse peptidyl- and aminoacyl-tRNAs. The column-eluted samples were concentrated to 20 µl by vacuum-evaporation. Subsequently, free SH-groups were blocked by addition of iodoacetamide to 0.1 M final concentration. Then an appropriate volume of 5 times concentrated 35 electrophoresis buffer (Schägger and von Jagow, (1987) was added

to the samples and they were heated to 70°C for 5 min and applied to the gel.

A gel system for SDS-gel electrophoresis (Schägger and von Jagow, 1987), resolving polypeptides between 1 and 100 kD, was 5 used to analyse the translation products. Gels (0.5 mm thick) were run at room temperature overnight at 100 V. Gels were quantified with Phoshor-Imager instrument (Molecular Dynamics). 10 ^{14}C -Rainbow low weight protein markers (Amersham, U.K.) from 2.2 to 46 kD were used to estimate the molecular weights of the gel 10 bands.

Band-shift assay

The binding of synthesised GR DBD to a target was analysed by a band shift assay (Dahlman-Write et al., 1990).

15

RESULTS

In the present work we have labelled the DNA recognising helix of the DNA binding domain (DBD) of the human glucocorticoid receptor (GR DBD) (Härd et al., (1990) and Dahlman-Wright et al., 20 (1990)). The DNA recognising helix is flanked by the first threonine (T) and the first glutamine (Q) in the GR DBD sequence (Fig 1). Translating ribosomes can therefore be made to pause at these flanking codons by blocking charging of either tRNA^{Thr} or tRNA^{Gln}. This was done by excluding Thr or Gln from the 25 translation mix and by using tRNA synthetase fractions defective either in ThrRS or GlnRS activity (Material and Methods).

A new technique to translate natural mRNAs, covalently linked to a column matrix, has been applied to ^{14}C labelling of amino acids T456-G470 (Fig 1) in the DNA binding domain (DBD) of the 30 glucocorticoid receptor (GR).

Three different translation mixes were prepared (Materials and Methods). The first mix, A, for translation of the N-terminal part of GR DBD contained all amino acids (unlabelled) except Thr and had a defective ThrRS. The second mix, B, for translation of 35 the T456-G470 region contained nine ^{14}C labelled amino acids, no Gln and was defective in GlnRS activity. The third mix, C,

contained all amino acids (unlabelled) and was deficient in release factor RF1 activity. This last mix was used to translate the C-terminal part of GR DBD.

Ribosomes (500 pmol) together with fmet-tRNA^{fMet} and initiation factors were applied to a column (0.4 ml) with mRNA for GR DBD coupled to it (Materials and Methods). Initiation of translation was obtained by 10 min incubation at room temperature.

Subsequently, translation mix A was applied to synthesise the fragment M397-L455 (Fig 1) of GR DBD.

After 5 min incubation with mix A the column was washed with PMGPT buffer (polymix buffer plus 1 mM GTP) and translation mix B was added for synthesis of the labelled peptide region. The column wash after this step was applied to a protein gel (fig 2), showing that some small peptide fragments had been eluted (lanes B1, B2) by the rinsing.

The complete GR DBD was synthesised by applying mix C. Fig 2 shows that the column wash after this step (lanes C1-C3) contained a small amount of GR DBD eluted from the column (lane 2) together with small amounts of short peptides. It was not until after RF1 addition to the column (lane R5), that the dominant fraction of region labelled GR DBD was eluted by washing (lanes R6 and R7). The column wash at this last step contained region labelled GR DBD together with RF1 and no other components. From this point a completely pure product could be obtained in a single gel filtration step separating GR DBD from RF1.

According to band shift assays (not shown) the in vitro synthesised region-labelled DBD was fully active in DNA binding.

To get almost pure region-labelled product (fig 2) it was necessary to eliminate the RF1 activity of mix C. This is illustrated in Fig 4, which shows an experiment where the last translation mix, C, contained a synthetase fraction with fully active RF1. In this experiment we also used ³⁵S Cys of high specific activity in mixes A, B, and C to make drop-off of peptides visible in the gel (fig 3). In this case the full length GR DBD eluted from the column after translation with mix C (lanes C1 and C2), together with all components of that mix. Some S-S

bonds were formed between free ^{35}S Cys and SH-groups of proteins in the translation mixes, so that EF-Tu, which is the main component of the translation mixes, could be seen in the different lanes of fig 3. The absence of any full-length GR DBD 5 product in lanes A1-A3 demonstrates that during translation with mix A the ribosomes are indeed stalled at the first Gln codon.

The amount of region-labelled GR DBD produced in these experiments (figs 2,3) was about 2 pmol, with about 140 pmol of 10 column coupled mRNA (Materials and Methods).

The amount of synthesised region-labelled GR DBD increased almost linearly with the amount of ribosomes applied to the column in the initiation step (fig 4). This experiment demonstrates that it is the efficiency of ribosome initiation, 15 rather than the amount of available column-coupled mRNA, that limits the protein yield.

COMMENTS

One advantage of the solid phase technique in the present 20 context is that it greatly facilitates switching between labelled and unlabelled amino acids during translation. An alternative method to prepare region labelled proteins is to translate mRNA in solution and then to purify pausing ribosomes with the nascent peptide by gel filtration. This, more cumbersome, method has also 25 been used to produce region labelled GR DBD (not shown, gel filtration on Sephacryl® S300 HR, Pharmacia Biotech AB). Accordingly the labelling aspect of the invention can be generalised to be carried out with or without the use of solid phases.

30

LEGENDS TO THE FIGURES

Fig. 1 The amino acid sequence (SEQ ID NO 4) of the glucocorticoid receptor (GR) DNA binding domain (DBD) with the description of the labelling procedure; numbering refer to the 35 rat GR.

Fig. 2 Synthesis of ^{14}C -region labelled GR DBD. Elongation was started by adding 0.15 ml of Translation Mix A (TMA) to the column. After 3 min of elongation TMA was removed from the column by washing 5 times with 0.6 ml PMGTP (polymix buffer plus 1 mM GTP). Then 0.15 ml of Translation Mix B, containing nine ^{14}C -labelled amino acids was added and elongation continued for 5 minutes. After this, 0.45 ml of PMGTP was added and the column eluent was collected (sample B1). The column was subsequently washed four times with 0.6 ml of PMGTP buffer (samples B2-B5). Samples B1-B2 were applied to the gel in lanes B1-B2. Next, we added 0.25 ml of Translation Mix C, allowed 5 min of elongation before adding 0.25 ml of PMGTP after which the column eluent was collected as sample C1. The column was subsequently washed 4 times with 0.6 ml of PMGTP (samples C2-C5). Samples C1-C3 were applied to the gel in lanes C1-C3. Finally, 500 pmol of RF1 in 0.6 ml of PMGTP buffer was added to the column (sample and lane R5). The completed region-labelled GR DBD was eluted from the column by washing with two 0.6 ml volumes of PMGTP buffer (samples and lanes R6 and R7).

20

Fig. 3 Synthesis of region labelled GR DBD. Highly radioactive ^{35}S Cys was added to all translation mixtures to monitor possible drop-off from the ribosomes. Elongation was started by adding 0.15 ml of translation mix A (TMA) to the column. After 3 minutes of elongation 0.45 ml of PMGTP was added and the column eluent collected (sample A1). The column was washed four times with 0.6 ml PMGTP (samples A2-A5). Samples A1-A3 were applied to gel in lanes A1-A3. Lanes B1-B3 correspond to samples B1-B3 from the column after elongation with Translation Mix B (TMB) lacking charged Gln-tRNA isoacceptors. Lanes C1-C3 correspond to samples C1-C3 from the column after elongation with translation mix C (TMC) with native RF1. See the legend to Fig. 2 for more details.

Fig. 4 Dependence of amount of synthesised region labelled GR DBD upon amount of ribosomes used in the initiation step. Amounts

of pure GR DBD after each column was calculated from intensities of GR DBD lanes on gel (not shown).

1. REFERENCES:

2. Baranov V L et al., Meth. Enzymol. 59 (1979) 382-397
3. Bax A., Ann. Rev. Biochem. 58 (1989) 223-256
4. Belitsina N V et al. FEBS Lett 57 (1975) 262
5. Belitsina N V et al., Meth. Enzymol. 60 (1979) 745-759
6. Boutillon C et al., Eur. J. Biochem. 231 (1995) 166-180
7. Clore G M et al., Science 252 (1991) 1390-1399
8. Corner et al., J. Med. Biol. 261 (1996) 108-124
9. Dahlman-Wright K et al., J. Biol. Chem. 265 (1990) 14030-14035
10. Dubnoff J S et al., Meth. Enzymol. XX (1971) 248-261
11. Ehrenberg M et al., In: The Practical Approach Series. Ed. Spedding G. IRL Press, Oxford and Washington (1990) 101-129
12. Gold et al., Meth. Enzymol. 185 (1990) 89-91
13. Härd T et al., Science 249 (1990) 157-160
14. Jelenc P C et al., Proc. Natl. Acad. Sci. USA 76 (1979) 3174-3178
15. Jelenc P C., Anal. Biochem. 105 (1980) 369-374
16. Kelmers A D et al., Meth. Enzymol. 20 (1971) 3-9
17. Luisi B F et al., Nature 352 (1991) 572-575
18. Methods in Enzymology, vol 239, "Nuclear Magnetic Resonance, Part C", Academic Press, San Diego (1994)
19. Muchmore D C et al., Meth. Enzymol. 177 (1989) 44-73
20. Ollins P et al., J. Biol. Chem. 264 (1989) 16973-16976
21. Pavlov M Yu et al., Arch. Biochem. Biophys. 328 (1996) 9-16
22. Ramesh V et al., Eur. J. Biochem 225 (1994) 601-608
23. Schägger H et al., Anal. Biochem. 166 (1987) 368-379
24. Soffientini A et al., Prot. Expr. Purif. 5 (1994) 118-124
25. Seno T et al., Biochem. Biophys. Acta 169 (1968) 80-94
26. Wagner E-G H et al. Eur. J. Biochem. 122 (1982) 193-197
27. Wuthrich K., Science 243 (1989) 45-50
28. Zhang H et al. The solution structure of the trp repressor-operator DNA complex. J. Mol. Biol. 238 (1984) 592-614
29. Zang et al., J. Mol. Biol. 242 (1994) 614-618
30. Zengel J-M et al. J. Bacteriol. 129 (1977) 1320-1329

PATENT CLAIMS

1. An in vitro method for production of a protein (polypeptide), characterized in that the mRNA encoding the protein is translated when bound to a solid phase.

5

2. The method of claim 1, characterized in that the translation is paused at one or more predetermined codons in the mRNA while maintaining the nascent polypeptide chain bound in a tRNA-ribosome-mRNA complex to the solid phase.

10

3. The method of claim 2, characterized in

a. that pausing is taking place at a stop codon,
b. that, if desired, the translation mixture used for the

15

translation up to the stop codon is removed from the nascent polypeptide bound in the tRNA-ribosome-mRNA complex to the solid phase, and

c. that release factor is added so that the nascent polypeptide is released from the solid phase.

20 4. The method of anyone of claims 1-3, characterized in

a. that one of the predetermined codons at which pausing is taken place is an amino acid encoding codon or a stop codon each of which is internally located in the amino acid encoding region of the mRNA, and

25 b. that translation up to this predetermined codon is done with a first translation mix devoid of one or more components necessary to read through this codon thereby causing the translation to become stalled, and

c. that translation is restarted by a second translation mix containing the factors necessary to read through the codon, and

30 d. that said first and second translation mixes differ with respect to the labelling of the amino acids present therein.

5. The method of claim 4, **characterized** in that pausing and restarting is taking place at two or more predetermined internal codons and that the translation mixes used are such that

- 5 a. a subsequent mix permits restart of translation at a predetermined codon at which translation is paused for a preceding mix, and
- b. a preceding mix and a subsequent mix differ from each other with respect to one or more labelled amino acids

10 6. A labelled polypeptide/protein, **characterized** in that its amino acid sequence contains one or more regions of two or more labelled amino acid residues in sequence.

15 7. A labelled polypeptide/protein, **characterized** in that there is at least one amino acid that is occurring twice and that this amino acid differs in labelling at at least two positions.

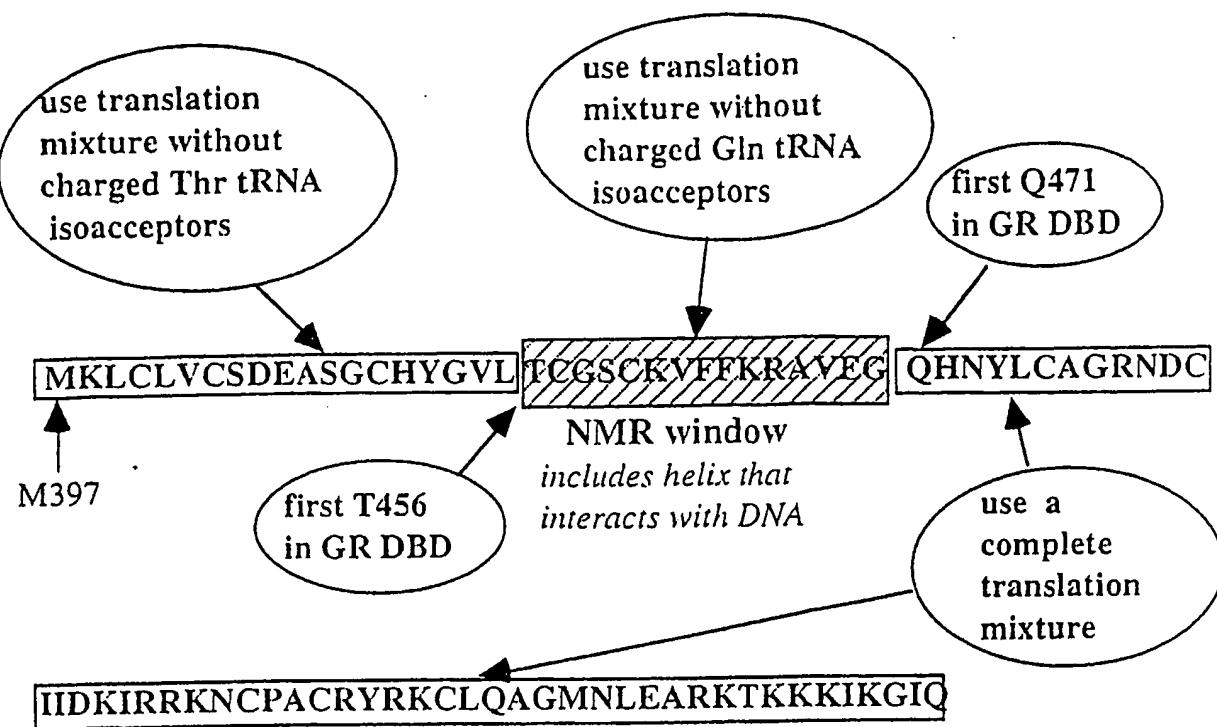


Fig 1

B1 B2 C1 C2 C3 R5 R6 R7 M



Fig. 2

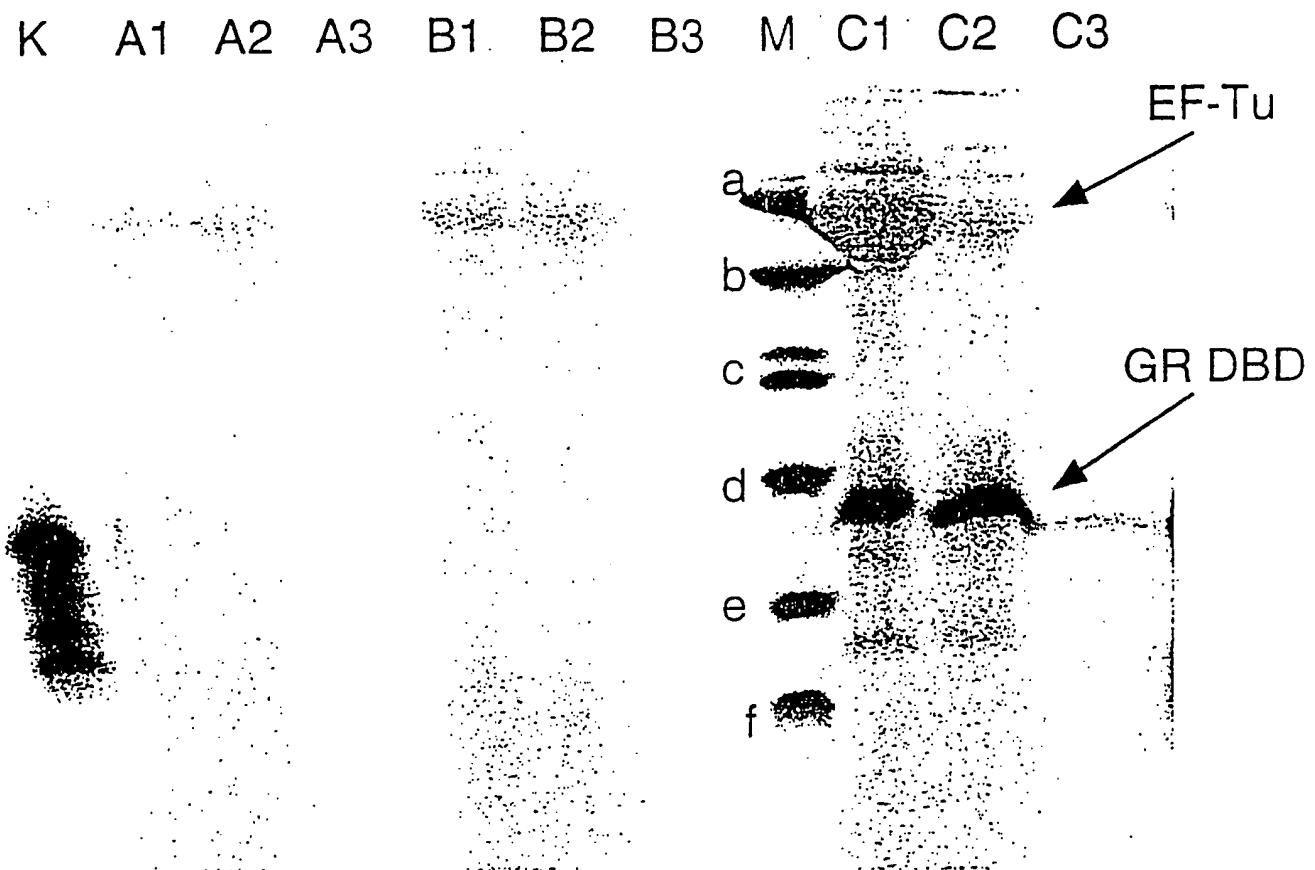


Fig 3

4 / 4

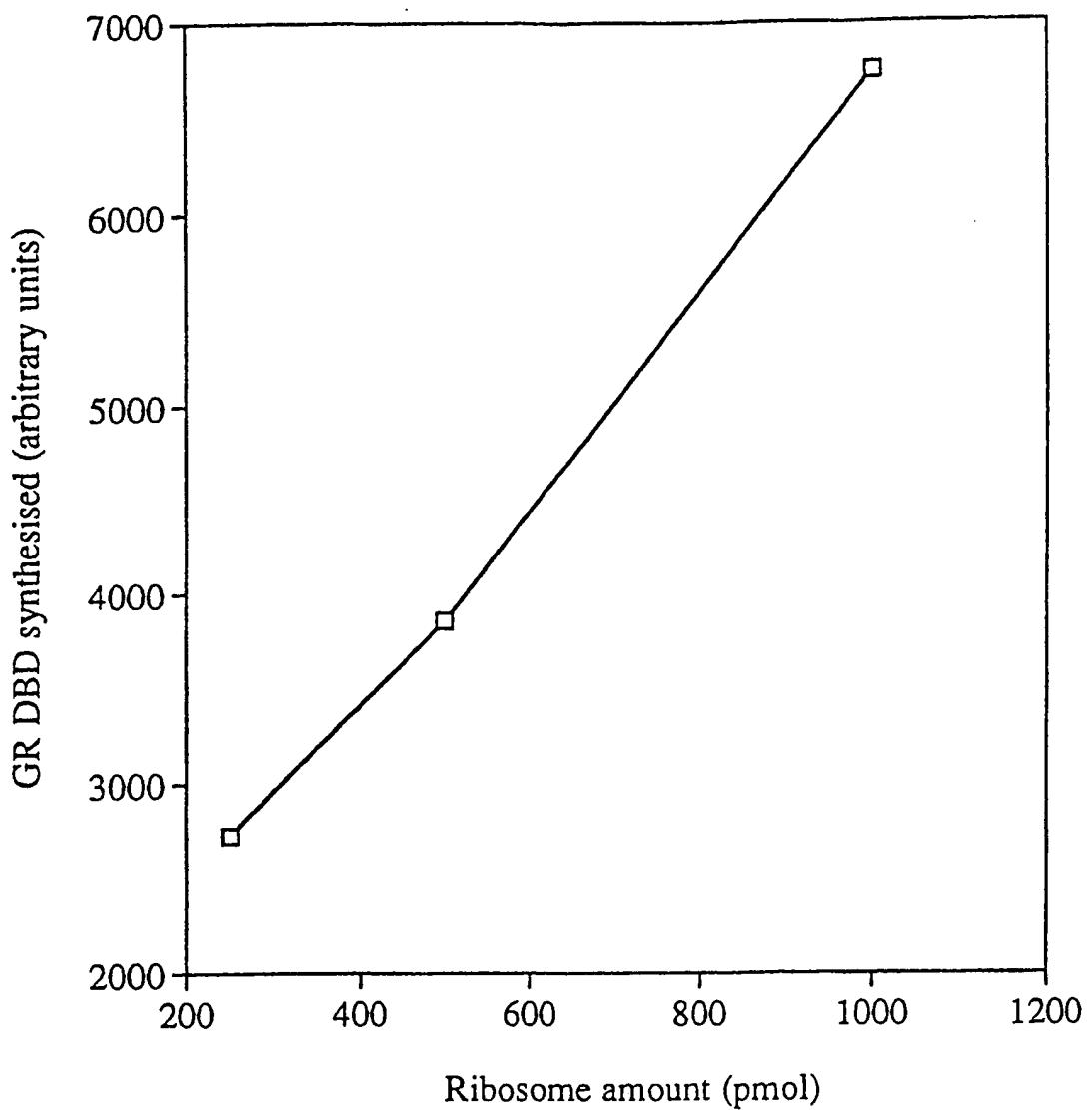


Fig 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE 97/01805

A. CLASSIFICATION OF SUBJECT MATTER

IPC6: C12P 21/00, C07K 1/04 // C07K 1/13
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC6: C12N, C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

MEDLINE, BIOSIS, CA, WPI

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	Dialog Information Services, file 154, MEDLINE, Dialog accession no. 09317798, Medline accession no. 98013345, Pavlov MY et al: "Synthesis of region-labelled proteins for NMR studies by in vitro translation of column-coupled mRNAs"; & Biochimie (FRANCE) Jul 1997, 79 (7) p415-22 --	1-5
X	Biotechnology and Bioengineering, Volume 37, 1991, Eiry Kobatake et al, "Translation of Immobilized Genetic Information by Yeast Cell-Free Protein Synthesizing System" page 723 - page 728	1
A	-- -----	2-5

 Further documents are listed in the continuation of Box C. See patent family, annex.

- * Special categories of cited documents:
- "A" document defining the general state of the art which is not considered to be of particular relevance
- "B" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

30 March 1998

02 -04- 1998

Name and mailing address of the ISA/

Swedish Patent Office

Authorized officer

INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE 97/01805

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see next sheet

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

1-5

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest



The additional search fees were accompanied by the applicant's protest.



No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE 97/01805

According to PCT rule 13.2, an international application shall relate to one invention only or a group of inventions linked by one or more of the same or corresponding "special technical features", i.e. features that define a contribution which each of the inventions makes over the prior art.

The present application claims an in vitro translation method for producing proteins. Claims 6 and 7 relates to product protection of labelled proteins. No unifying technical feature links the method of claim 1-5 and the proteins of claims 6-7. Accordingly, the following inventions were found:

Invention 1, claim 1-5: An in vitro translation method for producing proteins.

Invention 2, claims 6-7: Labelled polypeptides, at a further search this invention may be shown to be two unrelated inventions.

The search has been restricted to invention 1.